

CT248: Introduction to Modeling

Week #2: Classroom Challenge - Arrays

The goal of this exercise is to gain familiarity with manipulating arrays in MATLAB. The example builds an array of grades for students, adds aggregate values, and then calculates a frequency table. It uses a MATLAB function to generate data from an $N(0,1)$ distribution, and transforms this to grade averages.

(1) Create the grades matrix. Rows are individual data, columns are grade results.

```
clear;
N = 10;
ids = (1:N)';

mean1 = 55; sd1 = 5;
mean2 = 79; sd2 = 8;
mean3 = 46; sd3 = 15;
mean4 = 85; sd4 = 20;
mean5 = 67; sd5 = 2;
mean6 = 54; sd6 = 40;

rng(100);
sub1 = sd1 .* randn(N,1) + mean1;
sub2 = sd2 .* randn(N,1) + mean2;
sub3 = sd3 .* randn(N,1) + mean3;
sub4 = sd4 .* randn(N,1) + mean4;
sub5 = sd5 .* randn(N,1) + mean5;
sub6 = sd6 .* randn(N,1) + mean6;

% create results matrix
res = [sub1 sub2 sub3 sub4 sub5 sub6];
disp(['Initial simulated data...']);
disp(res);
```

Initial simulated data...

55.8047	71.6661	70.8806	62.3432	67.2666	50.2265
51.9244	70.7147	40.1486	78.5100	62.4884	60.1895
53.8050	77.2068	29.8380	58.2788	66.2023	24.8250
58.0749	100.6175	42.3620	92.6122	67.3995	26.5756
50.4930	83.1619	22.3194	125.2700	69.4458	-2.3720
56.7404	75.3223	32.5792	56.2800	66.3141	16.4034
59.7398	85.1279	53.7458	105.8411	66.7125	-10.1357
46.3506	72.5784	21.5531	89.8909	67.8241	54.3835
55.4856	76.9024	50.4526	112.7119	69.4578	37.5689
61.8929	91.6678	50.1755	89.9657	69.8418	58.8013

(2) Write the code to convert outliers to values within range, $> 100 \Rightarrow 100$. Note that a logical array/vector can be used as a subscript parameter to assign values to an array.

```
>> A = [1 2 3 4 5 6 7]

A =

     1     2     3     4     5     6     7

>> A(logical([1 0 0 0 0 0 1])) = 100

A =

    100     2     3     4     5     6    100
```

```
disp(res);
```

55.8047	71.6661	70.8806	62.3432	67.2666	50.2265
51.9244	70.7147	40.1486	78.5100	62.4884	60.1895
53.8050	77.2068	29.8380	58.2788	66.2023	24.8250
58.0749	100.0000	42.3620	92.6122	67.3995	26.5756
50.4930	83.1619	22.3194	100.0000	69.4458	-2.3720
56.7404	75.3223	32.5792	56.2800	66.3141	16.4034
59.7398	85.1279	53.7458	100.0000	66.7125	-10.1357
46.3506	72.5784	21.5531	89.8909	67.8241	54.3835
55.4856	76.9024	50.4526	100.0000	69.4578	37.5689
61.8929	91.6678	50.1755	89.9657	69.8418	58.8013

(3) Write the code to convert outliers to values within range, $< 0 \Rightarrow 0$

```
disp(res);
```

55.8047	71.6661	70.8806	62.3432	67.2666	50.2265
51.9244	70.7147	40.1486	78.5100	62.4884	60.1895
53.8050	77.2068	29.8380	58.2788	66.2023	24.8250
58.0749	100.0000	42.3620	92.6122	67.3995	26.5756
50.4930	83.1619	22.3194	100.0000	69.4458	0
56.7404	75.3223	32.5792	56.2800	66.3141	16.4034
59.7398	85.1279	53.7458	100.0000	66.7125	0
46.3506	72.5784	21.5531	89.8909	67.8241	54.3835
55.4856	76.9024	50.4526	100.0000	69.4578	37.5689
61.8929	91.6678	50.1755	89.9657	69.8418	58.8013

(4) Add the student IDs column

```
disp(res);
```

1.0000	55.8047	71.6661	70.8806	62.3432	67.2666	50.2265
2.0000	51.9244	70.7147	40.1486	78.5100	62.4884	60.1895
3.0000	53.8050	77.2068	29.8380	58.2788	66.2023	24.8250
4.0000	58.0749	100.0000	42.3620	92.6122	67.3995	26.5756
5.0000	50.4930	83.1619	22.3194	100.0000	69.4458	0
6.0000	56.7404	75.3223	32.5792	56.2800	66.3141	16.4034
7.0000	59.7398	85.1279	53.7458	100.0000	66.7125	0
8.0000	46.3506	72.5784	21.5531	89.8909	67.8241	54.3835
9.0000	55.4856	76.9024	50.4526	100.0000	69.4578	37.5689
10.0000	61.8929	91.6678	50.1755	89.9657	69.8418	58.8013

(5) Add columns for the grade average and overall score (columns 8 and 9).

Added columns for average and grade

1.0000	55.8047	71.6661	70.8806	62.3432	67.2666	50.2265	0	0
2.0000	51.9244	70.7147	40.1486	78.5100	62.4884	60.1895	0	0
3.0000	53.8050	77.2068	29.8380	58.2788	66.2023	24.8250	0	0
4.0000	58.0749	100.0000	42.3620	92.6122	67.3995	26.5756	0	0
5.0000	50.4930	83.1619	22.3194	100.0000	69.4458	0	0	0
6.0000	56.7404	75.3223	32.5792	56.2800	66.3141	16.4034	0	0
7.0000	59.7398	85.1279	53.7458	100.0000	66.7125	0	0	0
8.0000	46.3506	72.5784	21.5531	89.8909	67.8241	54.3835	0	0
9.0000	55.4856	76.9024	50.4526	100.0000	69.4578	37.5689	0	0
10.0000	61.8929	91.6678	50.1755	89.9657	69.8418	58.8013	0	0

(6) Add the mean mark in column 8 (using a loop)

res =

1.0000	55.8047	71.6661	70.8806	62.3432	67.2666	50.2265	63.0313	0
2.0000	51.9244	70.7147	40.1486	78.5100	62.4884	60.1895	60.6626	0
3.0000	53.8050	77.2068	29.8380	58.2788	66.2023	24.8250	51.6926	0
4.0000	58.0749	100.0000	42.3620	92.6122	67.3995	26.5756	64.5040	0
5.0000	50.4930	83.1619	22.3194	100.0000	69.4458	0	54.2367	0
6.0000	56.7404	75.3223	32.5792	56.2800	66.3141	16.4034	50.6066	0
7.0000	59.7398	85.1279	53.7458	100.0000	66.7125	0	60.8877	0
8.0000	46.3506	72.5784	21.5531	89.8909	67.8241	54.3835	58.7634	0
9.0000	55.4856	76.9024	50.4526	100.0000	69.4578	37.5689	64.9779	0
10.0000	61.8929	91.6678	50.1755	89.9657	69.8418	58.8013	70.3909	0

(7) Add overall grade (1, 2, 3, 4, or 5) to column 9.

Added the grade..

1.0000	55.8047	71.6661	70.8806	62.3432	67.2666	50.2265	63.0313	2.0000
2.0000	51.9244	70.7147	40.1486	78.5100	62.4884	60.1895	60.6626	2.0000
3.0000	53.8050	77.2068	29.8380	58.2788	66.2023	24.8250	51.6926	3.0000
4.0000	58.0749	100.0000	42.3620	92.6122	67.3995	26.5756	64.5040	2.0000
5.0000	50.4930	83.1619	22.3194	100.0000	69.4458	0	54.2367	3.0000
6.0000	56.7404	75.3223	32.5792	56.2800	66.3141	16.4034	50.6066	3.0000
7.0000	59.7398	85.1279	53.7458	100.0000	66.7125	0	60.8877	2.0000
8.0000	46.3506	72.5784	21.5531	89.8909	67.8241	54.3835	58.7634	3.0000
9.0000	55.4856	76.9024	50.4526	100.0000	69.4578	37.5689	64.9779	2.0000
10.0000	61.8929	91.6678	50.1755	89.9657	69.8418	58.8013	70.3909	1.0000

(8) Perform a frequency count on the grades

```
disp(['Frequency Table for Grades']);
disp(grades');
disp(freq);
```

Frequency Table for Grades

1 2 3

1 5 4

(9) Test on 1,000,000 rows.

Frequency Table for Grades

1 2 3 4 5

194393 504572 282101 18901 33

>>

>> sum(freq)

ans =

1000000